Available online at www.ejal.eu
http://dx.doi.org/10.32601/ejal.651341
Eurasian Journal of Applied Linguistics, 5(3), 473–493

EJAL
Eurasian Journal of Applied Linguistics

It was found that…: Introductory it Patterns by Native and Non-Native Authors

Fatih Güngör [a] *

[a] *Afyon Kocatepe University, Department of Foreign Language Education, Afyonkarahisar 03200, Turkey*

Received 22 April 2019 | Received in revised form 20 May 2019 | Accepted 26 June 2019

APA Citation:
Güngör, F. (2019). *It was found that*…: Introductory it patterns by native and non-native authors. *Eurasian Journal of Applied Linguistics, 5*(3), 473-493. **Doi: 10.32601/ejal.651341**

Abstract

Introductory *it* pattern, as in *it was found that*, is of significance in academic writing but the use of introductory *it* might be challenging especially for native- and non-native students and non-native academic writers. However, few studies have been conducted to compare the use of introductory *it* pattern by native and non-native scholars. This study investigates the frequencies, variability and functions of the introductory *it* patterns in the research articles of native and non-native academic professionals. The study uses data from the MCRA-L1 and MCRA-L2 corpora of MCRA (Multilingual Corpus of Research Articles) corpus. The size of each corpus was one million words. In order to extract introductory *it* patterns, four-word lexical bundles were searched for through WordSmith Tools with a cut-off point of 5 times per million words for 4-, 5- and 6-word bundles. The results revealed that there were 38 different introductory *it* patterns in the MCRA-L1 and 66 in the MCRA-L2, and the frequency and percentages showed the tendency of the Turkish authors to overuse the introductory *it*-structures in their research articles.

Keywords: academic writing; corpus study; research article; introductory-*it* pattern

1. Introduction

1.1. Background

In the past several decades since Swales (1990) revisited the term of genre, the analysis of academic discourse first became prominent as a domain of research in applied linguistics, and the structures and functions of academic discourse have been studied through contrastive rhetoric and corpus linguistics (Hyland, 2002). The reasons behind this enormous growth in the number of studies in academic discourse can directly be associated with the diversity of university students, the flourishing funding opportunities in teaching and learning, and the progression of English as the language of research and academia (Hyland, 2009). Also, the rise of academic discourse studies has compelled scholars from around the world to publish in English.

* Corresponding author. Tel.: +90-272-218-1740
E-mail address: fgungor@aku.edu.tr

Of academic discourses, research article is definitely "the pre-eminent genre of the academy" (Hyland, 2009, p. 67) and "the master narrative of our time" (Montgomery, 1996). Therefore, research article genre, compared to the student genres, has recently received a great deal of interest due to its distinctive purpose, audience, and rhetorical features (Hyland, 2008a, 2012) as a "norm-developing" practice (Swales, 1990, p. 31). The pedagogic concerns of English for Academic Purposes (EAP) and English for Specific Purposes (ESP) classrooms raised this interest for the discourse analysis and the scrutiny of linguistic features in research articles.

In this genre, on one hand, researchers are supposed to inform readers and report their findings objectively; on the other hand, they feel the need to persuade or convince readers by orienting them toward their viewpoint on propositional information through a range of rhetorical strategies (Charles, Hunston, & Pecorari, 2009; Hyland, 2008b; Hyland & Tse, 2005; Zhang, 2015). Then, metadiscourse has a vital role in the construction of knowledge through establishing a bridge between writers and readers who share the same or similar culture, academia and rhetorical practices in the same discourse community (Hu & Cao, 2011). That is to say, the use of metadiscourse strategically and aptly in academic writing can be associated with the acceptance potential of knowledge claims and the existence of an author in the discourse community involved (Hyland, 2005). Making knowledge claims is a primary goal of research articles (Basturkmen, 2009), and results and discussion sections as the placeholder of these claims deserve to be described and investigated thoroughly. Although the results section is the part that "drives the paper" (Cargill & O'Connor, 2006, p. 210) in the highly adopted IMRaD (Introduction-Method-Results-Discussion) model of the academia, it has received less attention than its counterparts (Williams, 1999).

Introductory *it* pattern constitutes an interesting phenomenon (Römer, 2009) as a frequent pattern in academic writing compared with other genres (Biber, Johansson, Leech, Conrad, & Finegan, 1999; Hewings & Hewings, 2002; Hyland, 2008a; Römer, 2009; Zhang, 2015). This pattern is used for "claiming objective necessity or certainty" (Collins, 1994, p. 20), giving "an appearance of objectivity and generality" (Herriman, 2000, p. 212), disguising personal and subjective evaluations of authors (Groom, 2005; Halliday, 1994; Herriman, 2000; Hewings & Hewings, 2002), and making writing impersonal and objective with passive matrix verbs (Collins, 1994; Zhang, 2015). This plethora of research reminded the statement of Hyland (2008, p. 3) that "academic writing is persuasive" and "at the heart of the academic persuasion is, then, writers' attempts to anticipate possible negative reactions to their claims". Although many recent studies revealed that introductory *it* pattern might be difficult and problematic for both native (Larsson, 2017) and non-native speakers (Hewings & Hewings, 2002, 2004; Hunston, 2002; Oakey, 2002) of English, most of them excluded introductory *it* patterns with passive matrix predicates (Zhang, 2015) and very little attention has been paid to investigating potential differences of introductory *it* patterns by native and non-native authors. Therefore, this study aims to compare the use of introductory *it* patterns in a corpus of research articles written by native speakers of English with

the use of introductory *it* patterns in a corpus of research articles written by non-native Turkish authors.

1.2. Review of literature

Introductory *it* pattern has attracted some attention in the literature, where it is commonly referred to as introductory *it* pattern (Groom, 2005; Larsson, 2017; Römer, 2009), subject extraposition (Biber et al., 1999), *it*-clauses (Hewings & Hewings, 2002), anticipatory it pattern (Ädel, 2014) and *it*-extraposition (Kaltenböck, 2005). In this study, introductory *it* pattern can be defined as a pattern which is composed of an introductory *it* and a nominal clause, and introductory *it* pattern should not have an anaphoric reference referring back to other pronouns in the text for its meaning (cf. Quirk, Greenbaum, Leech, & Svartvik, 1985). In other words, introductory *it* as "the structural requirement for an initial subject" seems to be void of meaning but it has a cataphoric reference to a clause in the latter part of the same sentence (Quirk et al., 1985, pp. 89, 349). Two examples were provided in (1) and (2) and the two subjects in each sentence were italicized for emphasis.

(1) *It* is important to recognize *that such freedom and flexibility are perhaps more possible for some schools and not others*. [MCRA-L1, SSE_2]

(2) *It* is clear *that the mentor is crucial to preservice teachers' development within the school.* [MCRA-L1, IES_40]

Nominal clauses can be classified into six main categories: *that*-clauses, subordinate interrogative clauses, subordinate exclamative clauses, nominal relative clauses, *to*-infinitive clauses, and –ing clauses (Quirk et al., 1985, pp. 1048-1049). The preceding "*for*" constructions were also included in *to*-infinitive clause category due to the note of Quirk et al. (1985, p. 1061) that "[t]he presence of a subject in a *to*-infinitive clause normally requires the presence of a preceding "*for*". Pronoun *it*, cleft-*it*, and prop *it* were excluded as the definition does not address these types of uses.

From EAP perspective, many researchers (e.g. Hewings & Hewings, 2002; Thompson, 2009) highlighted the significance of introductory *it* pattern in academic discourse studies. Although it is very common in academic writing (Biber, Johansson, Leech, Conrad, & Finegan, 1999; Hewings & Hewings, 2002; Hyland, 2008a; Römer, 2009; Zhang, 2015), native- and non-native students and non-native academic writers face problems with the use of introductory *it* patterns (Hewings & Hewings, 2002; Larsson, 2017; Rodman, 1991; Römer, 2009; Zhang, 2015). Introductory *it* pattern is ubiquitously used in academic writing in line with the principle that the information in a sentence is arranged from low to high value in English, which is called as end-focus or extraposition (Quirk et al., 1985). Writers tend to slide long and sophisticated information to the end of the sentence and to present new information there (Hyland & Tse, 2005), and this part, called as rheme by Halliday and Matthiessen (2004), constitute a message together with the first part, theme. However, some studies (Jacobs, 1995; Hewings & Hewings, 2002) prove that non-native writers might have

difficulty in using introductory *it* due to the absence of this structure in some languages.

To achieve anticipation and objectivity, writers tend to use introductory *it* patterns with passive matrix verbs (e.g. *it was found that*…) instead of the active structures (e.g. *I/we found that*…). This type of passive is closely related to the conventions of Anglo-American academic writing (Hinkel, 1997), and it emphasizes the process or experiment in research studies, not researcher (Hacker, 2003). The frequent use of passive verbs in academic writing (Hinkel, 2004; Quirk et al., 1985; Swales, 1990; Zhang, 2015) confirms the significance of tense, aspect and passive usage in academic writing (Hinkel, 2004; Swales & Feak, 2000). However, it is difficult and complex to use passive voice in English appropriately (Baratta, 2009; Hinkel, 1997) due to its contextual, lexical and semantic constraints (Hinkel, 2004; Jacobs, 1995). Furthermore, an overlap is evident between the studies on tense and voice uses in academic writing and the presentation of these features in academic writing classes (Hinkel, 2004). Although most writing instruction textbooks give present passive voice a minimal place and suggest it to be avoided in written discourse (Hinkel, 2004), there is no universal rule against use of passive voice unless people "have a good reason to do so" (Beason & Lester, 2012, p. 275). Considering the lack of studies on introductory *it*, especially the ones with passive matrix predicates (Zhang, 2015), a closer textual analysis seems to be necessary in academic writing.

2. Method

2.1. Data

This study used the data from the Multilingual Corpus of Research Articles (MCRA). MCRA is a three-million-word corpus that consists of published research articles (indexed in ERIC and SSCI) in different fields (e.g. mathematics teaching, science teaching, and language teaching etc.) of educational sciences. Güngör (2016) designed this "specialized corpus" (Kennedy, 1998, p. 20) including one-million-word corpus for each language variable, namely L1 English, L2 English and L1 Turkish. The size of each corpus is one-million-word because it can be said to be "large enough to adequately represent the occurrence of the features being studied" (Biber, 2006, p. 51).

Table 1. The corpus statistics

	MCRA-L1	MCRA-L2
Tokens (Running words)	1.000.019	1.000.009
The number of articles	165	206
Types (Distinct words)	25.445	24.743
Type/token ratio	2.61	2.58
STTR	38.13	33.35
STTR std. dev.	61.16	66.30
Sentences	34.821	34.978
Mean in words	26.47	27.46
Standard deviation	78.35	100.43

For the aims of this study, the two subcorpora of the MCRA (MCRA-L1 and MCRA-L2) were included in the study. The MCRA-L1 and the MCRA-L2 represents the English research articles written by native Anglo-American academic writers and native Turkish writers. Thus, the research articles in the MCRA-L1 were used as a reference to the ones in the MCRA-L2. Although some linguists (e.g. Jenkins, 2006) consider this a controversial issue against to the status of English as a Lingua Franca (ELF), using native speaker writing as a benchmark is a common practice (Granger, 2002) aimed to be helpful for EAP instruction due to the perception that "ENL [English as a Native Language] writing standards seem to be the only accepted norm" (Ingvarsdóttir & Arnbjörnsdóttir, 2013: 123). It has also been noted that many leading journals suggest academics to have their manuscripts checked by a native speaker (McKinley & Rose, 2018).

2.2. Data retrieval and processing

Since introductory *it* pattern and passive constructions in English largely consists of formulaic sequences (Hinkel, 2004; Wray, 2002), WordSmith Tools version 6.0 (Scott, 2016) was used to find 4-, 5-, and 6-word lexical bundles starting with introductory *it* pattern. The bundles were included in the study following the criteria mentioned in the introductory *it* pattern section. In addition to the linguistic criteria, the lexical bundles occurring at least five times per million words were included in the study for practical analysis of frequent bundles because being frequent in academic writing might mean that learners frequently will need those structures which are very beneficial in their writing (Gilquin, 2006). The number of the bundles retrieved from the MCRA-L1 and the MCRA-L2 were 77 and 276 respectively. Of six main categories of nominal clauses, two categories, namely *that*-clauses and *to*-infinitives were found in the corpus queries of the present study, and the preceding "*for*" constructions were also included in *to*-infinitive clause category due to the note of Quirk et al. (1985, 1061) that "[t]he presence of a subject in a *to*-infinitive clause normally requires the presence of a preceding "*for*". The overlapped lexical bundles (e.g. *it was also found that the* because of the article) were excluded to avoid the retrieval of inflated numbers of bundles. This exclusion yielded 38 and 66 lexical bundles for the MCRA-L1 and the MCRA-L2 respectively.

Table 2. The most frequent 5 lexical bundles in the MCRA-L1 and the MCRA-L2

Lexical Bundles in L1	Frequency	Lexical Bundles in L2	Frequency
It is important to	77	It was found that	193
It is clear that	31	It is seen that	165
It is possible that	31	It was determined that	118
It is difficult to	20	It was seen that	116
It is possible to	19	It was observed that	112

The basic frequency counts clearly show the reliance of non-native authors on the use of lexical bundles in their research articles. Staples, Egbert, Biber, and McClair (2013) also argue without distinguishing the functional or structural categories that L2 writers use many more lexical bundles rather than the native speakers do. Although the sizes of both corpora were the same, the number of the lexical bundles in the MCRA-L2 corpus was almost two times more than the number of the bundles in the MCRA-L1. The most frequent five bundles can be seen in Table 2. This table shows that the most frequent five bundles in the MCRA-L2 appeared at least 112 times per million words, and it reveals the tendency of non-native authors to use the lexical bundles frequently in their research articles.

2.3. A functional classification of introductory it patterns

The previous studies (e.g. Cortes, 2013) confirmed the link between lexical bundles and discourse functions. To scrutinize these functions, the current study used the taxonomies developed by Hewings and Hewings (2002) and Larsson (2017). Although Larsson (2017) excluded the attribution category from further analyses, it was not possible in this study due to the significant number of attributions in the MCRA-L2. Furthermore, the observation category was added to the classification of Hewings and Hewings (2002) due to the retrieval of this kind of bundles in the MCRA-L1 and the MCRA-L2.

Although many studies distinguished these functional categories on the basis of word semantics, it might be difficult to assign functions objectively (Larsson, 2017). Therefore, this study paid attention to semantics, linguistic features, and the context but supporting the assignment of functions with linguistic evidence was a priority for classification. Some of the lexical bundles were found to have more than one function as acknowledged by some studies (Biber, Conrad, & Cortes, 2004; Byrd & Coxhead, 2010; Salazar, 2010, 2014). In such situations, classifying bundles according to their most common use might be a good practice (Biber, Conrad, & Cortes, 2004) so the lexical bundles were categorized under the category in which they are most commonly used in the current study.

Hedges are known to be based on “plausible reasoning rather than certainty”, and researchers in social sciences and humanities use hedges twice as common than in hard sciences to eschew the direct involvement in the text (Hyland, 2011, p. 179). Also, hedges can be used to refrain from taking on full commitments to the second or extraposed part of the statement or mitigate the forcefulness of these commitments.

In other words, academics moderate the certainty of their statements and attain the fidelity of the audience for their claims (Hyland, 1994). Hedges are mostly conveyed through modal auxiliary verbs (e.g. may, might, and could), adjectival, adverbial and nominal modal expressions (e.g. possible, perhaps, and probability), and modal lexical verbs (e.g. believe and assume). Attitude markers are concerned with the writer's evaluation towards the content of the extraposed subject of a claim. This evaluation might be related to the worthwhileness of information (e.g. *it is worth pointing out*) or value judgment of the content (e.g. *it is important*). Emphatics (e.g. *should*, *must*, and *it is clear*) amplifies what author claim in the extraposed subject. Attributions, the fourth functional category, might be realized by ascribing a claim to a reference (e.g. *it is stated that*). Observations, the last category, aim to present the extraposed subject neutrally and do not have an interpersonal function.

3. Results & Discussion

3.1. An overall picture

Altogether, 2140 lexical bundles were identified as introductory *it* patterns in two corpora. Table 4 presents the number of function types and tokens for the introductory *it* patterns across two corpora, and the data were visualized in Figure 1 and Figure 2 for a graphical overview. The frequencies were not normalized as the sizes of each corpus were one-million words. The type and token frequencies of each function were compared through UCREL log-likelihood calculator across two corpora, and some noteworthy statistical differences were found (see Table 3). The asterisks showing the significance level were given under the table as a remark. The type numbers of two functional categories and the token numbers of four functional categories exhibited statistically significant differences. These results seem to point out the differences in the functions of the introductory *it* patterns between expert and novice writing, as a case reported by Larsson (2017) and Römer (2009).

Table 3. Function types and tokens

Functions	Type/Token	Hedges	Attitude markers	Emphatics	Attribution	Observation	Total
MCRA-L1	Type	13	18	3	0	4	38
	Token	652	993****	160	0	124	1929
MCRA-L2	Type	8	13	6	9***	30****	66
	Token	592	784	376****	740****	4292****	6784

*= p<0.05, **=p<0.01, ***=p<0.001, ****=p<0.0001

When the native Anglo-American academic writers were compared to the Turkish academic writers, certain functional differences were observed. While attitude markers, emphatics, attribution and observation categories exhibited statistically significant differences in terms of frequency of tokens, attribution and observation categories significantly differed in terms of frequency of types and tokens. In the subsequent subsections, the results were presented under the functional categories of introductory *it* patterns.

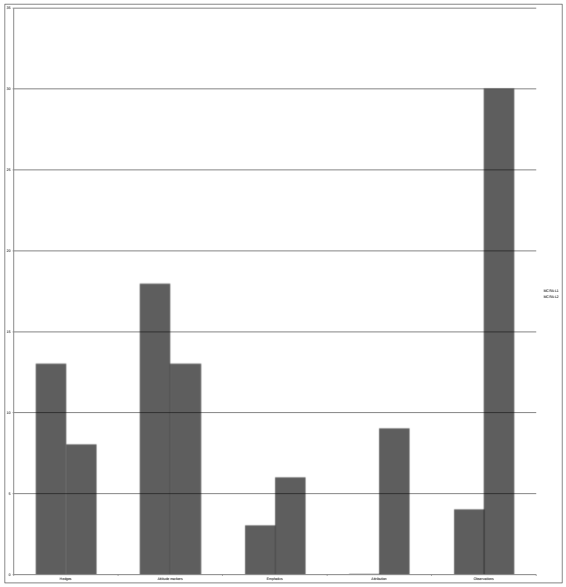

Figure 1. Frequencies of function types for the introductory *it* patterns

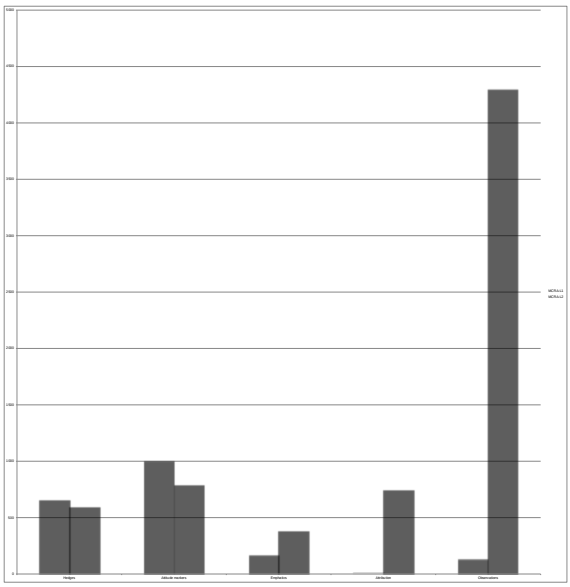

Figure 2. Frequencies of function tokens for the introductory *it* patterns

3.2. Attitude markers

Attitude markers were primarily formed as *be*+adjective with some modifications of the adjective, as in *it is evident that* or *it is clear that*. I was able to identify five out of six functions (including importance, validity, difficulty, expectancy and desirability excluding adequacy) occurring in the same pattern. Most of the bundles in the MCRA-L1 and the MCRA-L2 include the adjectives indicating importance, validity and difficulty in both corpora. It was also possible to see the instances of some expectancy and desirability meaning patterns in both corpora. The lack of adequacy meaning pattern might be related to the nature of academic writing as Römer (2009) also did not encounter the adequacy function in four apprentice and expert academic writing corpora including advanced student writings and research articles

Table 4. Attitude markers in MCRA-L1 and MCRA-L2

L1 English	it is apparent that (*f*=5), it is clear that (*f*=31), it is difficult to (*f*=20), it is important for (*f*=10), it is important that (*f*=11), it is important to (*f*=77), it is impossible to (*f*=5), it is interesting that (*f*=8), it is interesting to (*f*=5), it is not surprising that (*f*=7), it is reasonable to (*f*=7), it is useful to (*f*=6), it was clear that (*f*=11), it is also important to (*f*=7), it is essential that (*f*=7), it was important to (*f*=9), it is not possible to (*f*=9), it was not possible to (*f*=6)
L2 English	it is important to (*f*=52), it is clear that (*f*=21), it is important for (*f*=20), it is important that (*f*=17), it is obvious that (*f*=16), it is difficult to (*f*=13), it is emphasized that (*f*=10), it is essential to (*f*=10), it is remarkable that (*f*=9), it is evident that (*f*=8), it is reasonable to (*f*=8), it is essential that (*f*=7), it is impossible to (*f*=5)

As can be seen in the following examples, these expressions extrapose the second part of the sentence in line with the end-focus principle of English (Hyland & Tse, 2005; Quirk et al., 1985). When the extraposed part in a sentence is combined with validity adjective like “clear”, it refers to the writers’ endeavor to depersonalize their

opinions (Hyland & Tse, 2005) and show the validity of a statement (Halliday & Matthiessen, 2004). Biber et al. (1999) also confirms that the *it* v-link ADJ that pattern is noticeably in academic prose and closely linked with the validity meaning group.

(3) "…it is clear that teachers play a central role..." (MCRA-L1, SPE16)

(4) "Accordingly, it is clear that critical thinking will be an effective means..." (MCRA-L2, LE1)

The most frequent bundle was *it is important to* among attitude markers, in both corpora with modifications of different verbs. For instance, the bundles in the MCRA-L1 ended with the verbs such as note (*f*=20), recognize (*f*=8), understand (*f*=4), include (*f*=4) and acknowledge (*f*=4). As can be seen from the bundles, academic writers use the *it* v-link ADJ to-inf pattern to imply a critical, objective and rational voice (Groom, 2005). In other words, these verbs might indicate that the authors tend to highlight and strengthen the power of preceding adjectives. However, Turkish academic writers mostly used this pattern to highlight the gap for the study with verbs such as know (*f*=6), note (*f*=5) and investigate (*f*=5). The common verb *note* was used frequently to mark something of note, as found by Hewings and Hewings (2002). The bundle *it is important to* was the most frequent bundle in some other studies, and the frequencies of this bundle goes downward from research articles to doctoral dissertations and master theses (Jalali, Rasekh, & Rizi, 2009). The frequencies of the bundle in the MCRA-L1 (*f*=77) and the MCRA-L2 (*f*=52) show the same tendency, and this might be related to the diversity of the verbs used by native academic writers.

(5) "It is important to note that each study was initiated..." (MCRA-L1, PSE5)

(6) "Therefore, it is important to investigate the mental health..." (MCRA-L2, IDES124)

The validity meaning group followed the lead of importance meaning pattern. Three attitude markers were signifying validity in each corpus, and these lexical bundles were expressed by *it* be ADJ *that*-clause pattern, mainly associated with validity function (Biber et al., 1999, p. 675) and the obvious group of adjectives (Francis, Hunston, & Manning, 1998, p. 481). Although the frequencies were almost same (*f*=47 in the MCRA-L1, *f*=45 in the MCRA-L2), the only difference between native and non-native authors was the use of different adjectives. While the adjectives clear and apparent were the ones used in the MCRA-L1, three adjectives (clear, obvious and evident) appeared in the MCRA-L2. The examples of validity function can be seen in Examples (3) and (4).

The most frequent third meaning category was difficulty adjectives (e.g. difficult and impossible) for the lexical bundles in both corpora. The academic writers in both corpora seem to use the bundles with the adjective *difficult* as this adjective should be followed by an infinitive (Groom, 2005; Rodman, 1991).

(7) "Additionally, since it is impossible to determine…" (MCRA-L1, ITE4)

(8) "It is difficult to implement experiments related to..." (MCRA-L2, SE11)

The lexical bundles with desirability meaning consisted of it *be* ADJ *to*-inf structure, as stated by Biber et al. (1999, p. 721). The bundle *it is reasonable to* was observed in both corpora, and the bundle *it is useful to* was distinctive to the MCRA-L1. The following excerpts showed that these desirability patterns had a role to signal "the writer's evaluation of how desirable the proposition/phenomenon under discussion is" (Wang, 2018, p. 15). As the last function category, the expectation patterns (it *is interesting that, it is interesting to*, and *it is not surprising that*) were more diverse (n=3) and frequent (f=20) in the MCRA-L1. The only bundle with the expectation function pattern was *it is remarkable that*. The adjectives in the expectation patterns (surprising, interesting and remarkable) were assigned to the interesting and surprising group (expressed as expectation here) by Francis, Hunston and Manning (1998, pp. 483-484).

(9) "It is reasonable to suppose that language learning experience may be an important part…" (MCRA-L1, LE36)

(10) "…it is remarkable that participants are generally undergraduate students…" (MCRA-L2, IDES128)

Attitude markers used in the MCRA-L1 and the MCRA-L2 seem to be similar and semantically acceptable. These similar uses of the introductory *it* patterns as attitude markers confirm the impression of Hewings and Hewings (2002) that the differences were largely quantitative rather than qualitative in terms of attitude markers. Considering the aforementioned frequencies, it even seems that there were not major quantitative differences.

3.3. Emphatics

Of the introductory *it* functions, emphatics are the least frequent function in the MCRA-L1 and the MCRA-L2. The introductory *it* patterns in this category can be analyzed in two main structural patterns: *it is necessary to*+verb (12 times in the MCRA-L1 and 50 times in the MCRA-L2) and *it*+(modal)+passive verb (21 times in the MCRA-L1 and 39 times in the MCRA-L2). The structural difference between the two corpora was that the authors in the MCRA-L2 used passive structures instead of modal plus passive verbs. This difference seems to be related to the need of the non-native authors to express the implications of their studies directly with the verbs such as suggest and recommend.

(11) "It is necessary that the administrators, teachers, educators, and parents should behave responsibly…" (MCRA-L2, IDES7)

(12) "Accordingly, it is suggested that science teachers implement student-centered instructional strategies…" (MCRA-L2, SE16)

(13) "Therefore, it is recommended that numbers in problems be compatible with children's cognitive development regarding numbers." (MCRA-L2, PSE3)

(14) "…it is necessary to explore the long-term implications of…" (MCRA-L1, PSE12)

(15) "…it is necessary to design a model that is sympathetic…" (MCRA-L1, EMSE3)

(16) "It should be noted that students in this study were asked to report characteristics…; as a consequence, they were unlikely to phrase their responses in the negative." (MCRA-L1, PE2)

Table 5. Emphatics in MCRA-L1 and MCRA-L2

L1 English	it is necessary to (*f*=12), it should also be noted that (*f*=7), it should be noted that (*f*=14)
L2 English	it is suggested that (*f*=19), it is suggested to (*f*=9), it is recommended that (*f*=6), it is necessary that (*f*=5), it was suggested that (*f*=5)

What most of the examples had already showed that emphatics were used to amplify what author suggested in the discussion and conclusion, implications or recommendation parts of research articles. In other words, the emphatics in both corpora were used to claim that their research has some essential outcomes for the stakeholders (e.g. teachers, students, parents and administrators) in educational sciences field. Introductory *it* patterns and emphatics as interrelated structures are typical characteristics of advanced learners' interlanguage since the frequent use of extraposition strengthens the emphatic and persuasive style of writing (Herriman, 2013). Hewings and Hewings (2002) also underline the similar use of this pattern in their dissertation and research article corpora. Some extra functions of the emphatics were observed in the MCRA-L1. For instance, Example 15 emphasized the rationale and value of the research to impress the reader, as also stated by other researchers (e.g. Hewings & Hewings, 2002; Hinkel, 2005). Example 16 called the readers' attention to the limitations of the research, and such emphatics were expressed to guide readers' attention on what they should note (Rodman, 1991).

3.4. Attribution

Although some other studies excluded the attribution category from the analysis, this study focused on the use of introductory *it* patterns including attribution. Eight attribution bundles were used 173 times in the MCRA-L2, and this excessive use of the attribution might show us the tendency of Turkish authors to use this structure. Considering the finding of Hewings and Hewings (2002) that student writers use the introductory *it* patterns including attribution 113% more than published writers, the use of this function in the MCRA-L2 seems to be parallel with the student writing.

Table 6. Attribution in MCRA-L1 and MCRA-L2

L1 English	-
L2 English	it is thought that (*f*=42), it is believed that (*f*=34), it is known that (*f*=23), it is considered that (*f*=21), it is stated that (*f*=19), it was reported that (*f*=14), it was stated that (*f*=14), it was thought that (*f*=12), it is claimed that (*f*=6)

Of the eight introductory *it* patterns, four (*it is known that, it was reported that, it was stated that*, and *it is claimed that*) gives specific attribution to the literature in most occurrences, and the other four (*it is thought that, it is believed that, it is considered that,* and *it is stated that*) includes general attribution with no such

references. In other words, most of the occurrences for this eight introductory *it* patterns referred to general attribution without a specific reference. However, Hyland (2004, p. 20) highlights that the lack of appropriate citation might hinder "distinguishing observation and presumption" and constructing an authorial self. The following examples show the use of introductory *it* patterns by the authors in the MCRA-L2:

(17) "It is thought that these students have little interest in historical subjects." (MCRA-L2, SSE11)

(18) "It is believed that this study will bring a remarkable contribution to the literature." (MCRA-L2, SSE23)

(19) "It is known that concept teaching is significant in physical sciences." (MCRA-L2, PSE17)

(20) "Therefore, it is considered that reading skills have important effects …" (MCRA-L2, EME11)

(21) "… it is stated that drama should be included in preschool education…" (Author, year; Author, year)." (MCRA-L2, IDES101)

(22) "… it was reported that science process skills of science teachers were not sufficient" (Author names). (MCRA-L2, SE74)

(23) "It is claimed that over five million children …" (Author, year). (MCRA-L2, SE3)

In addition to the attribution function, *it is thought that* was used to express the inferences of the authors as in Example 17. Example 18 shows that *it is believed that* points out the expectations of authors. Due to this kind of exceptions, classifying bundles according to their most common use might be a good practice (Biber, Conrad, & Cortes, 2004) so the lexical bundles were categorized under the category in which they are most commonly used in the current study.

3.5. Hedges

The comparison of the function types and tokens produced similar results for hedging devices. The authors in the MCRA-L1 overused the hedge function in terms of tokens and types. Although the difference was not statistically significant, this finding should be noted because the non-native students were statistically underused introductory *it* patterns to hedge claims in the previous studies (e.g. Hewings & Hewings, 2002; Hinkel, 2005; Larsson, 2017). In other words, the underuse might be interpreted as a characteristic of the academic writing of novice or student writers, and the authors in the MCRA-L1 seem to be progressing toward the academic writing style of their native counterparts. However, it should be cautiously interpreted because the similar uses except a few instances might show in Römer's (2009) study that native speaker status do not have a significant role for using hedges in both corpora.

Table 7. Hedges in MCRA-L1 and MCRA-L2

L1 English	it is hoped that (*f*=10), it was hoped that (*f*=6), it can be argued that (*f*=7), it can be seen that (*f*=8), it could be argued that (*f*=9), it is argued that (*f*=8), it is assumed that (*f*=7), it is likely that (*f*=17), it is possible for (*f*=5), it is possible that (*f*=31), it is possible to (*f*=19), it may be that (*f*=15), it would appear that (*f*=15)
L2 English	it is expected that (*f*=11), it was expected that (*f*=8), it is hoped that (*f*=5), it is possible to (*f*=63), it is possible to say that (*f*=21), it is assumed that (*f*=14), it is possible to state that (*f*=7), it is possible that (*f*=5)

(24) "It is hoped that this study will reveal Turkish content area teachers' in-class practices and beliefs…" (MCRA-L2, EMSE18)

(25) "Therefore, it is expected that self-efficacy plays an essential role…" (MCRA-L2, SE23)

(26) "It is possible to evaluate reading as behaviour." (MCRA-L2, IDES76)

(27) "It is assumed that this result arises from the fact that…" (MCRA-L2, ME15)

(28) "It is hoped that this information can assist others in the teacher preparation field…" (MCRA-L1, PSE21)

(29) "It can be argued that many children who had unhappy and unproductive school careers can point to teachers…" (MCRA-L1, SE17)

(30) "…it is assumed that social and environmental problems can best be resolved through…" (MCRA-L1, SSE5)

(31) "It is likely that these challenges are compounded for pre-service teachers…" (MCRA-L1, ME15)

(32) "…it is possible that these students were already accustomed to using chat messages..." (MCRA-L1, IT2)

(33) "It may be that these providers were less preoccupied with the issue of standards…" (MCRA-L1, SE17)

(34) "It would appear that collegial support and the role of the teacher educator are important facets…" (MCRA-L1, ME15)

As can be seen from Table 7, most of the hedges were passive structures. Master (1991) also observed that the passive structures in academic writing function as hedges. Another dominant structural category in both corpora was *it is* ADJ *that*/*to*/*for*. There were also some structures such as *it may be that* (See Example 33) and *it would appear that* (See Example 34) that were not realized in non-native corpora of some studies (e.g. Larsson, 2017) as in this study. Syntactic differences between non-native authors' L1 (Turkish) and their L2 (English) might play a role for the lack of these patterns; however, this claim is subject to confirmation with further studies.

3.6. Observation

The Turkish authors employed approximately 31 lexical bundles with the observation function and repeated these bundles 1081 times. This was far more than the occurrence of observation bundles in native published articles (four types repeated 31 times), and this noteworthy difference was also statistically significant at the level of $p<0.0001$. These results made it clear that the introductory *it* patterns function to depersonalize the claim as if the claim was an accepted opinion (Kaltenböck, 2005). This "implicit attribution of stance to the speaker/writer" (Biber et al. 1999, p. 977) reduces the responsibility of author and increases the objectivity by attributing the claim to an external authority.

Table 8. Observation in MCRA-L1 and MCRA-L2

L1 English	it is noted that (*f*=5), it was decided that (*f*=6), it was decided to (*f*=8), it was found that (*f*=12)
L2 English	it was found that (*f*=193), it is seen that (*f*=165), it was determined that (*f*=118), it was seen that (*f*=116), it was observed that (*f*=112), it is observed that (*f*=59), it was concluded that (*f*=46), it is determined that (*f*=30), it is understood that (*f*=29), it was revealed that (*f*=27), it was aimed to (*f*=19), it is aimed to (*f*=16), it is revealed that (*f*=14), it is found that (*f*=13), it was detected that (*f*=11), it is founded that (*f*=10), it was decided to (*f*=9), it was noticed that (*f*=9), it was also found out that (*f*=8), it was also observed that (*f*=8), it was noted that (*f*=8), it was understood that (*f*=7), it is concluded that (*f*=6), it was established that (*f*=6), it is established that (*f*=5), it is indicated that (*f*=5), it was founded that (*f*=5), it was identified that (*f*=5), it was indicated that (*f*=5), it was recognized that (*f*=5)

For the sake of objectivity and depersonalization, some grammatically wrong structures (e.g. *it is founded that* and *it was founded that*) were also found in the MCRA-L2 (see Example 35 and Example 36) but these mistakes were in two individual published papers. To see the disparity of lexical choices, if any, the most frequent reporting verbs in the MCRA-L2 were compared with the ones in the study of Hyland (2000, 2002), and the verb determine was found to be distinctive to the Turkish authors (see Example 37).

(35) "…it is founded that teachers were more satisfied with relationships with colleagues and manager…" (MCRA-L2, EMSE3)

(36) "…it was founded that two-factor of Gibson and Dembo's Teacher Efficacy Scale is not suitable…" (MCRA-L2, IDES97)

(37) "…it was determined that there were significant differences…" (MCRA-L2, PSE12)

The most frequent seven bundles in the observation category were found to be the distinctive bundles to Turkish authors in Güngör and Uysal's (2016) study, and these inferential bundles were used to signal inferences and conclusions from the data. Almost all of the bundles in the observation category were inferential bundles signaling the forthcoming results of the data. In this study, the bundle *it was found that* was shared by both corpora but the other six bundles were found to occur one time maximum per million words in Corpus of Contemporary American English (COCA) (Güngör & Uysal, 2016). The underuse of these inferential bundles in native corpora, COCA, might indicate the deviation from the norms of native academic writing with redundant and excessive use of clausal and passive structures. The aim

of this redundant and excessive use appears to disguise subjective nature of evaluations (Hewings & Hewings, 2002, 2004; Zhang, 2015).

(38) "…it was concluded that the jigsaw is effective…" (MCRA-L2, SE34)

(39) "…it was detected that there was no significant difference between…" (MCRA-L2, SE7)

(40) "It is understood that the principals' being emotionally deprived…cause their performances to be low." (MCRA-L2, IDES52)

(41) "…it is observed that 53 teachers from the 185 pre-school teachers…had special education student in their classes…" (MCRA-L2, PSE10)

(42) "…it was established that students … could solve algorithmic questions…" (MCRA-L2, ME8)

In addition to the inferential bundles, there were also some procedural bundles (*it was aimed to, it is aimed to*, and *it was decided to*) with the function to describe actions, processes, methods and activities. These bundles show that the doer of the action is clear but the authors in the MCRA-L2 attempt to depersonalize the process.

(43) "…it was aimed to determine the applicability of the simultaneous distance education method…" (MCRA-L2, ME10)

(44) "In the current study, it was decided to choose the students from a high school." (MCRA-L2, SSE11)

4. Conclusions

This study aimed to map out the functional differences of introductory *it* patterns in research articles of native and non-native authors. The results showed that the authors in the MCRA-L2 used the introductory *it* patterns more frequently (74% and 252% more in terms of types and tokens respectively) than the authors in the MCRA-L1, as in some other studies (e.g. Hewings & Hewings, 2002). However, this greater use did not disperse across the five functional categories. While the authors in the MCRA-L2 made greater use of introductory *it* patterns for indicating emphatics (100% more), attribution (800% more) and observation (675% more), they make less use of introductory *it* patterns for hedging (38% less) and marking attitudes (28% less). These results seem to support the findings of Larsson (2017) and Römer (2009) that native speaker status is a factor in the use of introductory *it* patterns. Furthermore, native- and non-native students and non-native academic writers were found to have problems with the use of introductory *it* patterns (Hewings & Hewings, 2002; Larsson, 2017; Rodman, 1991; Römer, 2009; Zhang, 2015).

The analysis of introductory *it* patterns and the overall results presented in this study suggest how the authors in the MCRA-L2 endeavor to convince readers of the truth of their propositions. Herriman (2013) also points out the link between introductory *it* patterns and persuasive style of writing. To achieve the persuasive style of academic writing (Hyland, 2008b), authors combine rheme and theme through

introductory *it* patterns. Thus, writers are able to build anticipation (Kaltenböck, 2003), disguise subjective nature of evaluations (Hewings & Hewings, 2002, 2004; Zhang, 2015) and attitude (Herriman, 2000), obscure the writer's identity (Charles, 2006), and appear more objective in evaluations (Herriman, 2000). In other words, this pattern is used for "claiming objective necessity or certainty" (Collins, 1994, p. 20), giving "an appearance of objectivity and generality" (Herriman, 2000, p. 212), disguising personal and subjective evaluations of authors (Groom, 2005; Halliday, 1994; Herriman, 2000; Hewings & Hewings, 2002), and making writing impersonal and objective with passive matrix verbs (Collins, 1994; Zhang, 2015). However, Publication Manual of APA (American Psychological Association, 2009, pp. 77, 67) suggest to "use the active rather than passive" to avoid dangling modifiers and to write concise papers by eliminating redundancy and overuse of passive for the economy of expression. This plethora of research reminded the statement of Hyland (2008, p. 3) that "academic writing is persuasive" and "at the heart of the academic persuasion is, then, writers' attempts to anticipate possible negative reactions to their claims". However, this extraposed and persuasive style of non-native writing might give a justification for grammatically correct but non-native sounding introductory *it* patterns (Herriman, 2013), and Lorenz (1998) names this as the interlanguage of advanced learners.

Although it needs to confirm with further studies on introductory *it* patterns, syntactic differences between L1 and L2 might play a role for the overuse of these patterns. Many researchers in global (e.g. Allen, 2010; Paquot, 2013, 2014; Pérez-Llantada, 2014) or Turkish (Bal, 2010; Güngör & Uysal, 2016; Öztürk, 2014) context pointed to the potential crosslinguistic influence. As a result of these claims, Güngör and Uysal (2016), in his crosslinguistic study, found out that the most frequent seven observation bundles in this study seems to be transferred from Turkish authors' native language (Turkish) and the bundles distinctive to the Turkish authors were clausal structures such as introductory *it* structures and passive verb fragments. Certainly, the crosslinguistic influence might not be the only factor affecting the use of the introductory *it* patterns. For instance, Işık-Taş (2018), in a recent study, provided evidence that the discipline, L1 orientation and publication context determined the rhetorical choices of Turkish authors. As the MCRA-L2 greatly consists of research articles from locally-oriented journals in Turkey, their choice to publish nationally might have influenced their linguistic choices in establishing authorial identity, as suggested by Işık-Taş (2018). However, this crosslinguistic study was limited to the personal pronouns. Therefore, it can be suggested to investigate a small-scale crosslinguistic corpus for a detailed scrutiny of introductory *it* patterns with the aim of the syllabus design for the course of English for academic purposes tailored to Turkish authors.

As for the practical implication, academic programs at universities might offer academic English for research purposes courses especially for postgraduate students and these courses should be served by the experts having a reputable research career in the relevant discipline. In addition to these tailored courses, universities within the

English as a foreign language setting might establish academic writing centers to help researchers but this should not be limited to just translation and/or proofreading. These centers can extend their services to structuring manuscripts, proofreading and editing with the presence of author(s), referencing, critical writing, organizing academic workshops, and providing one-to-one tutorials. Academic writing centers can organize some workshops and teach researchers even to make basic queries on Corpus of Contemporary American English and British National Corpus for the development of their manuscripts. To conclude, universities and academic writing centers are expected to support researchers to become autonomous writers and to encourage them to be a part of their writing adventure.

Acknowledgements

This study was funded by the Scientific Research Unit of Afyon Kocatepe University (Project No: 17.KARİYER.122). I also thank for the valuable feedback of anonymous reviewers.

References

Ädel, A. (2014). Selecting quantitative data for qualitative analysis: A case study connecting a lexicogrammatical pattern to rhetorical moves. *Journal of English for Academic Purposes, 16*, 68-80. https://doi.org/10.1016/j.jeap.2014.09.001

Allen, D. (2010). Lexical bundles in learner writing: An analysis of formulaic language in the ALESS Learner Corpus. *Komaba Journal of English Education, 1*, 105-127. Retrieved from http://park.itc.u-tokyo.ac.jp/eigo/KJEE/001/105-127.pdf

American Psychological Association. (2009). *Publication manual of American Psychological Association* (6th ed.). Washington, DC: American Psychological Association.

Bal, B. (2010). *Analysis of four-word lexical bundles in published research articles written by Turkish authors* (Unpublished MA dissertation). Georgia State University, Atlanta, GA. Retrieved from https://scholarworks.gsu.edu/alesl_theses/2/

Baratta, A. M. (2009). Revealing stance through passive voice. *Journal of Pragmatics, 41*(7), 1406-1421. https://doi.org/10.1016/j.pragma.2008.09.010

Basturkmen, H. (2009). Commenting on results in published research articles and masters dissertations in language teaching. *Journal of English for Academic Purposes, 8*(4), 241-251. https://doi.org/10.1016/j.jeap.2009.07.001

Beason, L. & Lester, M. (2012). *A commonsense guide to grammar and usage* (6th ed.). Boston: Bedford/ St. Martin's.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers* (Vol. 23). Amsterdam: John Benjamins Publishing.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at ...: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*(3), 371–405. https://doi.org/10.1093/applin/25.3.371

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Longman.

Byrd, P., & Coxhead, A. (2010). On the other hand: Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney Papers in TESOL, 5*(5), 31-64. Retrieved from http://faculty.edfac.usyd.edu.au/projects/usp_in_tesol/pdf/volume05/Article02.pdf

Cargill, M., & O'Connor, P. (2006). Developing Chinese scientists' skills for publishing in English: Evaluating collaborating-colleague workshops based on genre analysis. *Journal of English for Academic Purposes, 5*(3), 207-221. https://doi.org/10.1016/j.jeap.2006.07.002

Charles, M. (2006). Revealing and obscuring the writer's identity: Evidence from a corpus of theses. In R. Kiely, P. Rea-Dickins, H. Woodfield & G. Clibbon (Eds.), *Language, culture and identity in applied linguistics* (pp. 147-161). London: BAAL/Equinox.

Charles, M., Hunston, S., & Pecorari, D. (2009). *Academic writing: At the interface of corpus and discourse*. London: Continuum International Publishing.

Collins, P. (1994). Extraposition in English. *Functions of Language, 1*(1), 7–24. https://doi.org/10.1075/fol.1.1.03col

Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes, 12*(1), 33-43. https://doi.org/10.1016/j.jeap.2012.11.002

Francis, G., Hunston, S., & Manning, E. (1998). *Grammar Patterns 2: Nouns and adjectives*. London: HarperCollins.

Gilquin, G. (2006). Highly polysemous words in foreign language teaching: How to give learners a flying start. In Proceedings of *the 7th Conference on Teaching and Language Corpora*, Université Paris 7 – Denis Diderot, 1-4 July 2006 (pp. 58-60). Retrieved from http://hdl.handle.net/2078.1/75870

Granger, S. (2002). A birds-eye view of learner corpus research. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition, and foreign language teaching* (pp. 3-33). Amsterdam: John Benjamins

Groom, N. (2005). Pattern and meaning across genres and disciplines: An exploratory study. *Journal of English for Academic Purposes, 4*(3), 257-277. https://doi.org/10.1016/j.jeap.2005.03.002

Güngör (2016). *Crosslinguistic analysis of lexical bundles in L1 English, L2 English, and L1 Turkish research articles* (Unpublished doctoral dissertation). Graduate School of Education, Gazi University, Ankara.

Güngör, F., & Uysal, H. H. (2016). A comparative analysis of lexical bundles used by native and non-native scholars. *English Language Teaching, 9*(6), 176-188. http://dx.doi.org/10.5539/elt.v9n6p176

Hacker, D. (2003). *A writer's reference* (5th ed.). Boston: Bedford/St Martin's.

Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). London: Foreign Language Teaching & Research Press.

Halliday, M. A. K., & Matthiessen, C. (2004). *An introduction to functional grammar* (3rd ed.). London: Arnold.

Herriman, J. (2000). The functions of extraposition in English texts. *Functions of Language, 7*, 203-230. https://doi.org/10.1075/fol.7.2.03her

Herriman, J. (2013). The extraposition of clausal subjects in English and Swedish. In K. Aijmer & B. Altenberg (Eds.), *Advances in corpus-based contrastive linguistics: Studies in honor of Stig Johansson* (pp. 233-260). Amsterdam: John Benjamins Publishing.

Hewings, M., & Hewings, A. (2002). "It is interesting to note that…": a comparative study of anticipatory 'it' in student and published writing. *English for Specific Purposes, 21*(4), 367-383. https://doi.org/10.1016/S0889-4906(01)00016-3

Hewings, M. & Hewings, A. (2004). Impersonalizing stance: A study of anticipatory 'it' in student and published academic writing. In C. Coffin, A. Hewings & K. O'Halloran (Eds.), *Applying English grammar: Functional and corpus approaches* (pp. 101–116). London: Arnold.

Hinkel, E. (1997). Indirectness in L1 and L2 academic writing. *Journal of Pragmatics, 27*(3), 361-386. https://doi.org/10.1016/S0378-2166(96)00040-9

Hinkel, E. (2004). Tense, aspect and the passive voice in L1 and L2 academic texts. *Language Teaching Research, 8*(1), 5-29. https://doi.org/10.1191/1362168804lr132oa

Hinkel, E. (2005). Hedging, inflating, and persuading in L2 academic writing. *Applied Language Learning, 15*(1/2), 29-53. Retrieved from http://www.dliflc.edu/wp-content/uploads/2014/04/ALL15-1.pdf

Hunston, S. (2002). Pattern grammar, language teaching, and linguistic variation: Applications of a corpus-driven grammar. In R. Reppen, S. M. Fitzmaurice & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 167-183). Amsterdam/Philadelphia: John Benjamins.

Hu, G., & Cao, F. (2011). Hedging and boosting in abstracts of applied linguistics articles: A comparative study of English-and Chinese-medium journals. *Journal of Pragmatics, 43*(11), 2795-2809. https://doi.org/10.1016/j.pragma.2011.04.007

Hyland, K. (1994). Hedging in academic writing and EAP textbooks. *English for Specific Purposes, 13*(3), 239–256. https://doi.org/10.1016/0889-4906(94)90004-3

Hyland, K. (2000). *Disciplinary discourses: Social interaction in academic writing*. London: Longman Pearson Education.

Hyland, K. (2002). Activity and evaluation: Reporting practices in academic writing. In J. Flowerdew (Ed.), *Academic discourse* (pp. 115-130). London: Routledge.

Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing*. USA: University of Michigan Press.

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse studies, 7*(2), 173-192. https://doi.org/10.1177/1461445605050365

Hyland, K. (2008a). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics, 18*(1), 41-62. https://doi.org/10.1111/j.1473-4192.2008.00178.x

Hyland, K. (2008b). Persuasion, interaction and the construction of knowledge: representing self and others in research writing. *International Journal of English Studies, 8*, 1-23. Retrieved from https://revistas.um.es/ijes/article/view/49151

Hyland, K. (2009). *Academic discourse: English in a global context*. London: Continuum.

Hyland, K (2011). Academic discourse. In K. Hyland & B. Paltridge (Eds.), *The continuum companion to discourse analysis* (pp. 171-184). London: Continuum.

Hyland, K., & Tse, P. (2005). Hooking the reader: A corpus study of evaluative that in abstracts. *English for Specific Purposes, 24*(2), 123-139. https://doi.org/10.1016/j.esp.2004.02.002

Ingvarsdóttir, H., & Arnbjörnsdóttir, B. (2013). ELF and academic writing: A perspective from the expanding circle. *Journal of English as a Lingua Franca, 2*(1), 123–145. https://doi.org/10.1515/jelf-2013-0006

Işık-Taş, E. E. (2018). Authorial identity in Turkish language and English language research articles in Sociology: The role of publication context in academic writers' discourse choices. *English for Specific Purposes, 49*, 26-38. https://doi.org/10.1016/j.esp.2017.10.003

Jacobs, R. A. (1995). *English syntax*: *A grammar for English language professionals*. Oxford: Oxford University Press.

Jalali, H., Rasekh, A. E., & Rizi, M. T. (2009). Anticipatory 'it' lexical bundles: A comparative study of student and published writing in applied linguistics. *Iranian Journal of Language Studies, 3*(2), 177-194. Retrieved from http://www.ijls.net/pages/volume/vol3no2.html

Jenkins, J. (2006). Current perspectives on teaching World Englishes and English as a Lingua Franca. *TESOL Quarterly, 40*, 157-181. https://doi.org/10.2307/40264515

Kaltenböck, G. (2003). On the syntactic and semantic status of anticipatory it. *English Language & Linguistics, 7*(2), 235-255. https://doi.org/10.1017/S1360674303001096

Kaltenböck, G. (2005). It-extraposition in English: A functional view. *International Journal of Corpus Linguistics, 10*(2), 119-159. https://doi.org/10.1075/ijcl.10.2.02kal

Kennedy, G. (1998). *An introduction to corpus linguistics*. London: Longman.

Larsson, T. (2017). A functional classification of the introductory *it* pattern: Investigating academic writing by non-native-speaker and native-speaker students. *English for Specific Purposes, 48*, 57-70. https://doi.org/10.1016/j.esp.2017.06.001

Lorenz, G. (1998). Overstatement in advanced learners' writing: stylistic aspects of adjective intensification. In S. Granger & G. Leech, *Learner English on computer* (pp. 53-66). New York: Taylor & Francis.

Master, P. (1991). Active verbs with inanimate subjects in scientific prose. *English for Specific Purposes, 10*(1), 15-33. https://doi.org/10.1016/0889-4906(91)90013-M

McKinley, J., & Rose, H. (2018). Conceptualizations of language errors, standards, norms and nativeness in English for research publication purposes: An analysis of journal submission guidelines. *Journal of Second Language Writing, 42*, 1-11. https://doi.org/10.1016/j.jslw.2018.07.003

Montgomery, S. (1996). *The scientific voice*. New York: The Guildford Press.

Oakey, D. (2002). Formulaic language in English academic writing. In R. Reppen, S. M. Fitzmaurice, & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 111-130). Amsterdam: John Benjamins Publishing Company.

Öztürk, Y. (2014). *Lexical bundle use of Turkish and native English writers: A corpus- based study* (Unpublished MA dissertation). Anadolu University, Eskişehir, Turkey.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. New York, NY: Longman.

Paquot, M. (2013). Lexical bundles and L1 transfer effects. *International Journal of Corpus Linguistics, 18*(3), 391-417. https://doi.org/10.1075/ijcl.18.3.06paq

Paquot, M. (2014). Cross-linguistic influence and formulaic language: Recurrent word sequences in French learner writing. *EUROSLA Yearbook, 14*(1), 240-261. https://doi.org/10.1075/eurosla.14.10paq

Pérez-Llantada, C. (2014). Formulaic language in L1 and L2 expert academic writing: Convergent and divergent usage. *Journal of English for Academic Purposes, 14*, 84-94. https://doi.org/10.1016/j.jeap.2014.01.002

Rodman, L. (1991). Anticipatory *it* in scientific discourse. *Journal of Technical Writing and Communication, 21*(1), 17-27. https://doi.org/10.2190/PWJ6-AG95-MCQD-RG1W

Römer, U. (2009). The inseparability of lexis and grammar. *Annual Review of Cognitive Linguistics, 7*, 141–163. https://doi.org/10.1075/arcl.7.06rom

Salazar, D. (2010). Lexical bundles in Philippine and British scientific English. *Philippine Journal of Linguistics, 41*, 94–109. Retrieved from http://116.50.242.171/PSSC/index.php/pjl01/article/view/2561

Salazar, D. (2014). *Lexical bundles in native and non-native scientific writing: Applying a corpus-based study to language teaching* (Vol. 65). Amsterdam: John Benjamins Publishing Company.

Scott, M. (2016). *WordSmith Tools version 6*. Stroud: Lexical Analysis Software.

Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes, 12*(3), 214-225. https://doi.org/10.1016/j.jeap.2013.05.002

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

Swales, J. & Feak, C. (2000). *English in today's research world*. Ann Arbor: The University of Michigan Press.

Thompson, P. (2009). Shared disciplinary norms and individual traits in the writing of British undergraduates. In M. Gotti (Ed.), *Commonality and individuality in academic discourse* (pp. 53-82). Bern: Peter Lang.

Wang, Y. (2018). As Hill seems to suggest: Variability in formulaic sequences with interpersonal functions in L1 novice and expert academic writing. *Journal of English for Academic Purposes, 33*, 12-23. https://doi.org/10.1016/j.jeap.2018.01.003

Williams, I. A. (1999). Results sections of medical research articles: Analysis of rhetorical categories for pedagogical purposes. *English for Specific Purposes, 18*(4), 347-366. https://doi.org/10.1016/S0889-4906(98)00003-9

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Zhang, G. (2015). It is suggested that… or it is better to…? Forms and meanings of subject it-extraposition in academic and popular writing. *Journal of English for Academic Purposes, 20*, 1-13. https://doi.org/10.1016/j.jeap.2015.02.004
